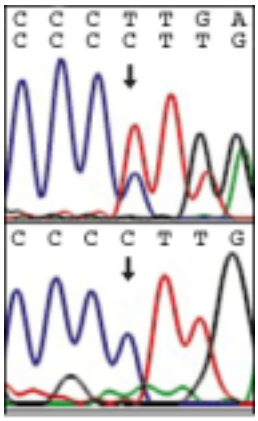
NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION
NCBI Coffee Break

6 August 2001

Article reference: CB21.060801
Coffee Break archives

nature REVIEWS Story by Jane Alfred, *Nature Reviews Genetics*

Opening the flood gates?



DNA sequence electropherograms. A cytosine insertion in the *NOD2* gene leads to a premature truncation of the protein product.

Click on the figure for more information.

It's been a long time coming, but now two papers report a clear cut identification by linkage mapping of a gene involved in a common human disorder Crohn's disease (CD). Importantly, they also indicate how the innate immune system might be involved in the aetiology of CD, because the identified gene *NOD2* encodes an intracellular receptor for bacterial lipopolysaccharides (LPS) that activates NFB, a target of the innate immune signalling pathway and a transcriptional regulator of inflammatory genes.

CD is a chronic inflammatory gut disorder, thought to be caused by an abnormal inflammatory response to enteric microbes. In 1996, a CD susceptibility locus, *IBD1* , was identified on chromosome 16. Little progress has been made since, but it is this locus that the two research teams one European, the other US-based tackled in their studies, using positional-cloning and candidate-gene strategies, respectively.

Hugot *et al.* took a decisive step when they identified association of CD to an allele of a chromosome-16 microsatellite marker. Despite the borderline significance of this association, the authors went on to identify putative transcripts in the region of this marker, and identified over 30 single nucleotide polymorphisms (SNPs) by sequencing the region from affected and unaffected individuals. Several turned out to be non-synonymous variants in a chromosome-16 gene, *NOD2* . Three of these SNPs each independently associated with disease susceptibility altered the leucine-rich repeat (LRR) region of NOD2, which is required for LPS recognition.

Having previously identified *NOD2* , Ogura *et al.* considered it a candidate for CD because of its chromosome-16 location. On sequencing the gene from CD individuals, they identified an insertion that caused two frameshift mutations in the LRR region and the premature truncation of NOD2. In *in vitro* assays, this mutant NOD2 produced considerably diminished levels of NFB activation in response to bacterial LPS compared to wild-type NOD2.

So how could *NOD2* contribute to susceptibility to CD? The innate immune system regulates the immediate immune response to bacterial pathogens, components of which are recognized in host immune cells by specific receptors, such as NOD2. A defect in this recognition might lead to an exaggerated inflammatory reaction being mediated by the adaptive immune system. Alternatively, NOD2 might act to trigger cytokines that dampen inflammatory responses. Although NOD2 does not account for all susceptibility to CD, it does provide a first glimpse into the aetiology of the disease and should speed the discovery of other CD loci and future therapies, and improve its diagnosis. These papers are hopefully the first of many such successes in grappling with the genetic basis of multifactorial, common disease.

Live PubMed searches
|| *NOD2* gene || *IBD1* gene || REVIEWS ||

MapViewer tutorial
|| Are there additional polymorphisms in the *NOD2* gene? || VIEW ANALYSIS ||

Further resources
|| *NOD2* in LocusLink|| *IBD1* in LocusLink || Genes and Disease || Medline Plus ||

Opening the flood gates?

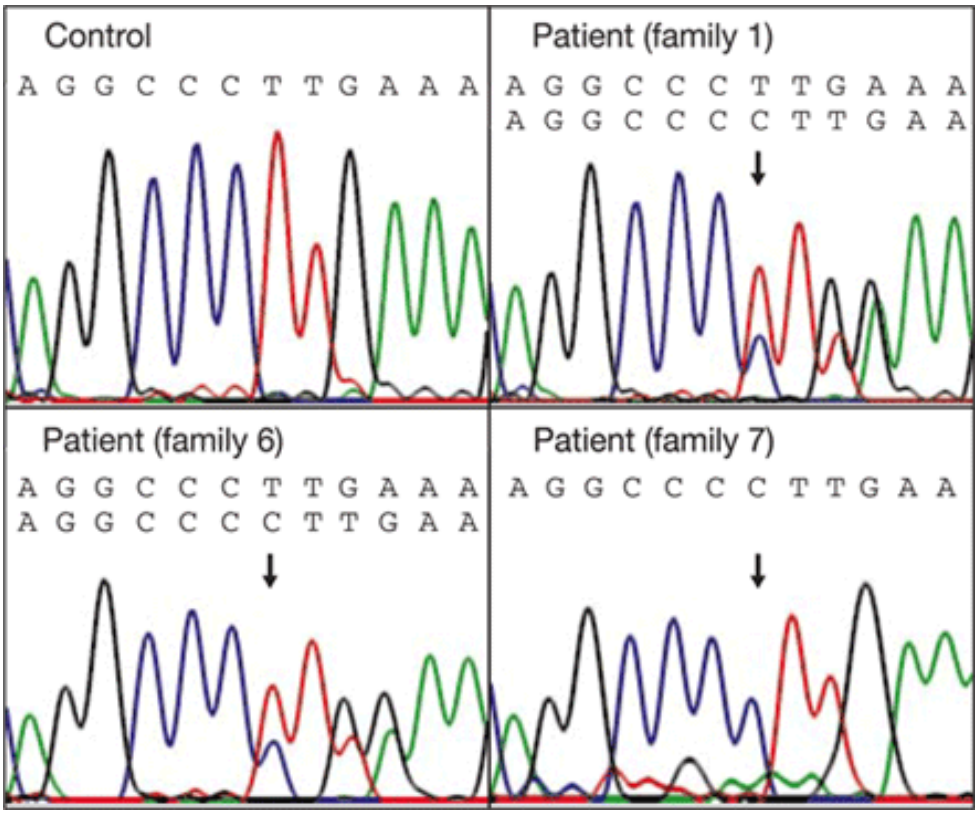


Figure 1. DNA sequence electropherograms of the *NOD2* gene. A portion of *NOD2* exon 11 DNA sequence from control and three CD-affected individuals. The control sequence codes for full-length NOD2 protein. The patients from families 1 and 6 are heterozygous for a cytosine insertion at position 3020 in the *NOD2* gene. The wild-type sequence in these panels is in the upper position and is read GCC-CTT-GAA. The sequence containing the cytosine insert is in the lower position and is read GCC-CCT-TGA. The extra cytosine base (marked by the arrows) causes a framshift mutation to occur, and the TGA sequence immediately downstream is recognized as a stop codon, causing the NOD2 protein to be truncated. The patient from family 7 is homozygous for the same cytosine insertion.

Reprinted by permission from Nature (Ogura Y, *et al.* (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 411(6837):603-6. [PubMed])

NCBI Tutorial Coffee Break

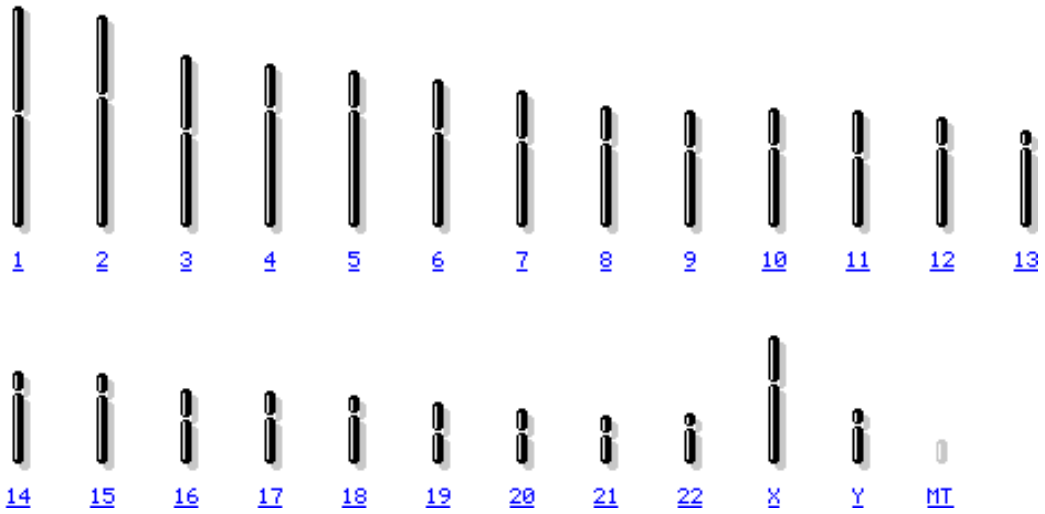
PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Help

Search for nod2 on chromosome(s) Find

Show linked entries Help FTP Advanced search

MapViewer facilitates gene searches of the human genome. Search terms are entered in the "Search for" box above. To search for the *NOD2* gene, click on the **Find** button above (marked by the red arrow).
Note: selecting other links will take you out of this tutorial.

Homo sapiens **genome view** **build 24** **BLAST search** the human genome



The NCBI Map Viewer presents a graphical view of the available Human Genome sequence data as well as cytogenetic, genetic, physical, and radiation hybrid maps.

The Map Viewer provides displays of the Human Genome sequence for the NCBI contigs (the 'Contig' map; see assembly description), the BAC tiling path (the 'GenBank' map), and the location of genes, STSs, FISH mapped clones, and variation on the contig sequence.

You can find genes or markers of interest by submitting a query against the whole genome, or a chromosome at a time. Results are indicated both graphically, as tick marks on the ideogram, and in a tabular format. The results table includes links to a chromosome graphical view where the gene or marker can be seen in the context of additional data. Alternatively, you can browse a chromosome by clicking on a chromosome link in the ideogram above.

Additional information on display control, navigation, and zoom control for the MapViewer is available in the **Help document**; descriptions of the **human maps** displayed are also provided. A separate document provides more detail about the status of the **human genome sequence data**.

NCBI Tutorial Coffee Break

PubMed | Nucleotide | Protein | Genome | Structure | PopSet | Taxonomy | OMIM | Help

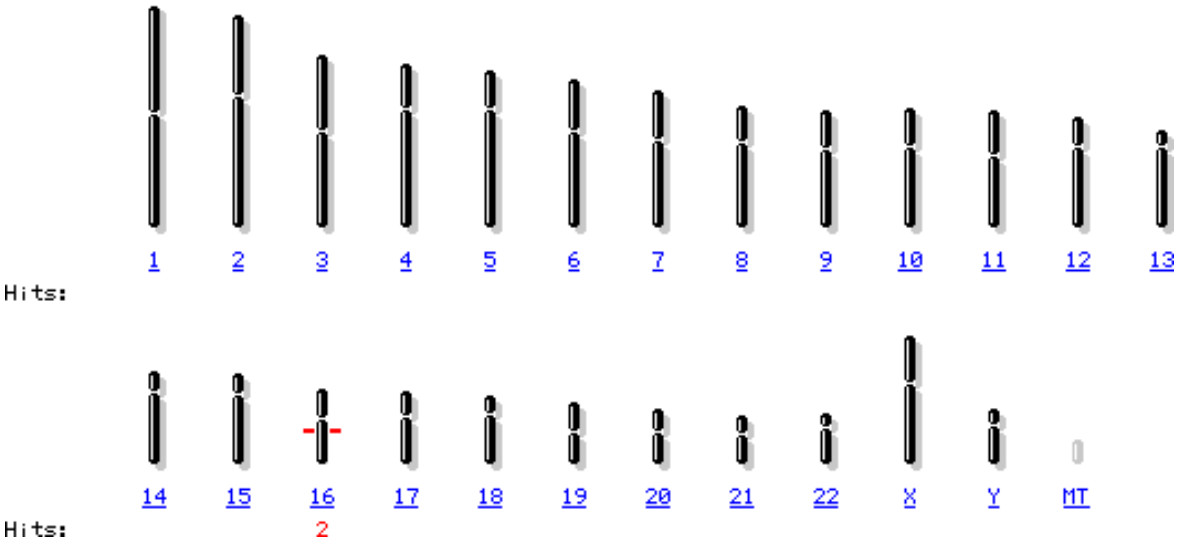
Search for nod2 on chromosome(s) Find

Show linked entries | Help | FTP | Advanced search

The search for *NOD2* has returned 2 matches, their positions marked by the red notches on chromosome 16. To view the *NOD2* gene within its genomic context, click on the "Genes seq" link marked by the red arrow at the bottom of the page.

Homo sapiens genome view build 24

BLAST search the human genome



Search results for query "nod2": 2 hits

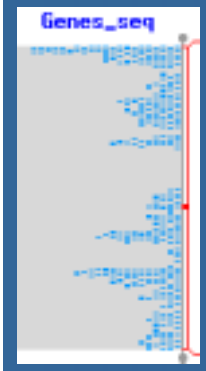
Chr	Match	Map element	Type	Maps
16	NOD2	605956	mim	Morbid
16	NOD2	NOD2	locus	Genes_cyto | Genes_seq

Map Viewer Help
Human Maps Help
FTP
Chr. 16 Resource

Data As Table View

Region Shown:
Go

out zoom in



Homo sapiens Map View build 24

Map Viewer displays the *NOD2* gene highlighted in red within the context of its neighbors on the 'Genes on Sequence Map' (because the Genes_seq link was selected in the previous page). To change the maps displayed below, click on the "Display settings" link marked by the red arrow.

Chromosome: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 **[16]** 17 18 19 20 21 22 X Y

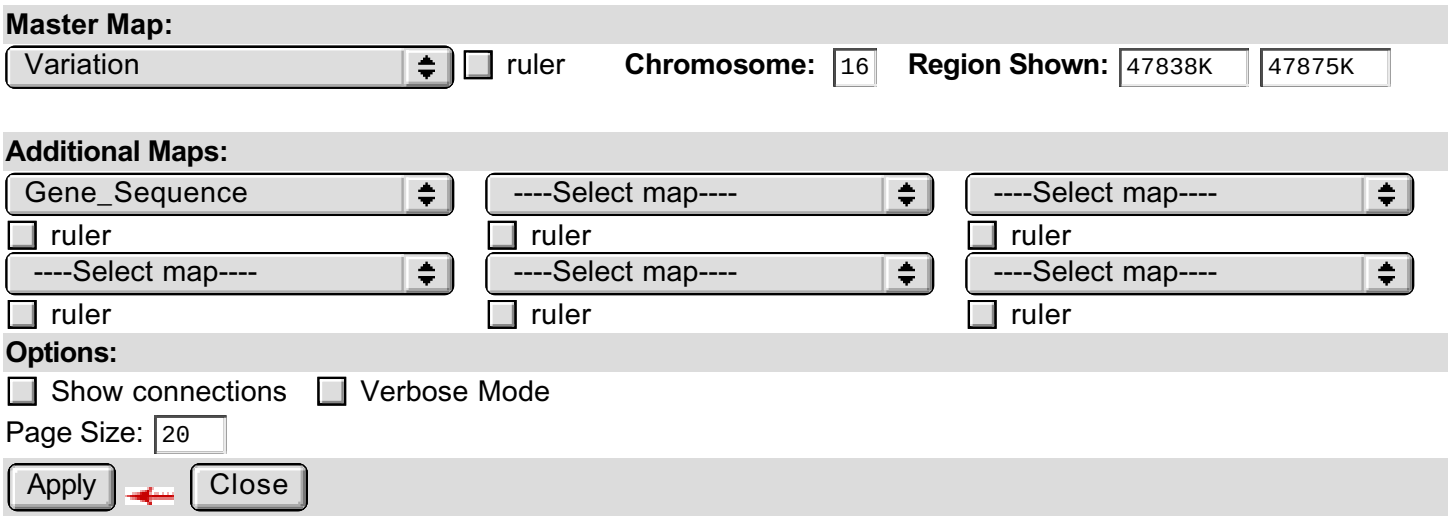
Query: nod2 [clear]

Master: Genes On Sequence Map

Display settings

Total Genes On Chromosome: **717** [7 not localized]
Region Displayed: **0-90M bp** Download/View Sequence
Genes Labeled: **20** Total Genes in Region: **710**

symb	orient.	links		cyto.	full name
SSTR5	-	sv	ev	16p13.3	somatostatin receptor 5
MGC13114	+	sv	ev	16	hypothetical protein MGC13114
CAB56184	+	sv	ev	16p13.3	hypothetical protein CAB56184
KIAA1306	+	sv	ev	16	KIAA1306 protein
FN14	+	sv	ev	16p13.3	type I transmembrane protein Fn14
KIAA0420	+	sv	ev	16	KIAA0420 gene product
LOC51283	-	sv	ev	16p13.13-16p12.3	apoptosis regulator
TNP2	-	sv	ev	16p13.13	transition protein 2 (during histone to protamine replacement)
KIAA0370	-	sv	ev	16	KIAA0370 protein
STX4A	+	sv	ev	16p13.13-16p12.3	syntaxin 4A (placental)
NOD2	+	sv	ev	16q12	NOD2 protein
SALL1	-	sv	ev	16q12.1	sal (Drosophila)-like 1
POLR2C	+	sv	ev	16q13-q21	polymerase (RNA) II (DNA directed) polypeptide C (33kD)
CDH5	+	sv	ev	16q22.1	cadherin 5, type 2, VE-cadherin (vascular epithelium)
HSD11B2	+	sv	ev	16q22	hydroxysteroid (11-beta) dehydrogenase 2
FLJ22593	-	sv	ev	16	hypothetical protein FLJ22593
KIAA0931	-	sv	ev	16	KIAA0931 protein
KIAA0703	+	sv	ev	16q22.1-q22.3	KIAA0703 gene product
ICSBP1	+	sv	ev	16	interferon consensus sequence binding protein 1
PCOLN3	-	sv	ev	16q24.3	procollagen (type III) N-endopeptidase

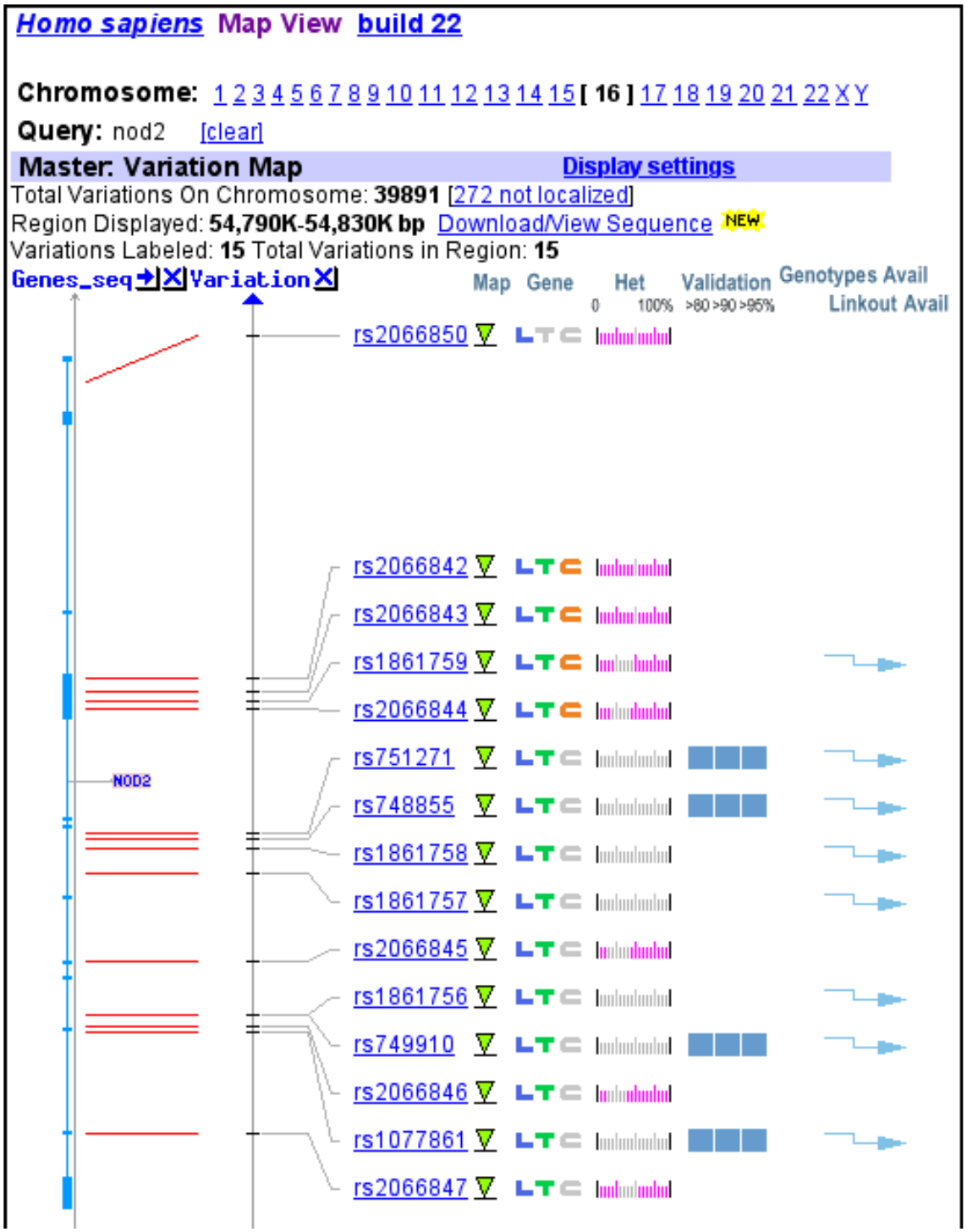


Selecting the "Display settings" link on a MapViewer page brings up this pop-up window control panel. For this tutorial, the Variation map has been selected as the master map, Gene_Sequence has been selected as an additional map, and the region shown has been set to 47838K-47875K on chromosome 16. Click the "Apply" button marked by the red arrow to continue.

Analysis Coffee Break

Opening the flood gates?

NOD2 **gene in the human genome**



MapViewer is an NCBI resource that can be used to display integrated information about the human genome. This information may be in the form of ideograms of the chromosomes that compose a genome, various integrated maps of individual chromosomes, or sequence data for specific regions. The map above depicts single nucleotide polymorphisms (SNPs) that have been found in the *NOD2* gene. The "Genes_seq" map to the far left is a graphical representation of the *NOD2* gene thick blue sections represent exons and thin blue sections represent introns. From this, it can be seen that *NOD2* is organized into 12 exons. The positions of reported *NOD2* SNPs are marked by the red lines.

The column labeled "Gene" is a graphical summary of several properties of the reported SNP markers. The symbol "L" shows that the marker is positioned within the locus, "T" designates markers that are positioned on the mRNA (although they may fall on an intron and not expressed in the protein), and "C" indicates markers that are found within the coding sequence (CDS) of a gene. On the variation map above, the SNP cluster IDs (preceded by 'rs' which stands for reference sequence) are hyperlinks to individual SNP records. SNP database records contain information such as whether or not a SNP represents a synonymous or non-synonymous change [1]. Of the SNPs reported above, their association with Crohn's disease if any remains to be determined.

[1] A synonymous change means that the SNP does not effect the amino acid coded for in the protein. However, a non-synonymous change results in a triplet that codes for a different amino acid, the occurence of which may have functional or structural implications for the protein.

Click this to close the window.